International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-12, October 2019

Statistical Relationship between Acceptance Rate and First Decision Time of Some Indexed Journals: A Correlation-Regression Technique

Pavan Kumar, W. Sridhar

Abstract: ***In this paper, we propose a correlation-regression approach to establish a statistical relationship between acceptance rate & first decision time of some indexed journals. We collect the data from two web sources: Elsevier journal finder, and Springer journal suggester. In this problem, we concentrate on the data of the acceptance rate and first decision time. To examine the relationship between these measures, we apply a statistical approach, which is based on correlation and regression analysis. We determine the relative error (RE) of the data collected. We plot the scatter diagram between the two measures. Karl Pearson correlation coefficient (CC) value is also calculated. All this analysis demonstrates that there is a moderate positive correlation between acceptance rate & first decision time.***

Keywords: ***Acceptance rate, first decision time, Regression, correlation.***

I. INTRODUCTION

Now a-days, the statistical approaches are very important tool for data analysis and comparison. In statistical approaches, the correlation and regression analysis are very popular. Zeger, S.L., (1988) proposed a regression model for time series of counts. Hair, J.F., Anderson, R.E., Tatham, R.L., & Black, W.C. (1998) introduced some regression models in their work: Multivariate data analysis. Natalie, M., & James, R. (2002) contemplated an improved batch means procedure for simulation output analysis. Koenker, R. (2004) studied a quantile regression for longitudinal data. Canay, I.A. (2011) formulated a simple approach to quantile regression for panel data. Azagba, S., Sharaf, M. F. (2012) studied the fruit and vegetable consumption and body mass index by applying a quantile regression approach.

Wang, J. (2012) presented a bayesian quantile regression mathematical problem for parametric nonlinear mixed effects models. Ismail, Azman et. al. (2013) studied and examined the mathematical relationship between service quality and customer satisfaction by applying a factor specific technique. Osman, Z., & Sentosa, I. (2013) studied the mediating effect of customer satisfaction on service quality and customer loyalty relationship.

Revised Manuscript Received on October 30, 2019.
***** Correspondence Author
Pavan Kumar*, Department of Mathematics, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India. Email: pavankmaths@gmail.com,
W. Sridhar, Department of Mathematics, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh , India. Email: sridharwuriti@gmail.com

They collected the real data from Malaysia Rural Tourism. Qiu, Heting; Li, & Xuewei, (2013) presented a study of combined evaluation of suppliers based on linear correlation for ungrouped data.

Bottai, M., Frongillo, E. A., & Sui, X. (2014) studied some uses of the quantile regression model to investigate the longitudinal association or mathematical relationship between physical activity and body mass index.

Wang, Yin-ying; Lin et. al. (2015) presented a study consists of the canonical correlation analysis between the fitness of organizational condition and integration possibility by applying the evidence of configurational methodology. Geraci, M. (2016) studied the estimation of regression quantiles in complex surveys with data missing at random by presenting an application to birthweight determinants. Coulom, Jean Shenai, & Vijay, (2018) studied the effect of alternative measures of distance on the correlation of real effective exchange rate returns. The approach which was used by these researchers is related to correlation analysis.

Some researchers considered the probabilistic as well as fuzzy type uncertainty. Recently, Kumar, P., & Keerthika, P. S., (2018) studied an inventory model with variable holding cost and partial backlogging under interval uncertainty by using global criteria method. Afterwards, Prameela, K.U., & Kumar, P., (2019) proposed the execution proportions of multi server queuing model with pentagonal fuzzy number by introducing an application of DSW algorithm approach.

The motivation behind this work is to develop a mathematical formula for studying the acceptance rate and the number of days giving the first decision for reputed journals. To the best of our knowledge, there is no such mathematical formula in literature. To fill this gap, the contribution of the present work is justified.

In the present work, we study the relationship between the journal measures consisting of acceptance rate and first decision time. We discuss some basic concepts in section 2. In following section 3, we describe some notations. The proposed problem definition is presented in section 4. In the section 5, we give some numerical calculations to calculate the correlation coefficient, mean relative error and the straight line equation. In the last, we conclude the work in the section 6.

II. BASIC CONCEPTS

Karl Pearson Correlation Coefficient (CC):

Correlation analysis is an attempt to measure the strength of relationships between two variables.

This is measured by means of a single number, which is called a correlation coefficient (CC).

CC generally measures the strength of the linear association between x and y. Karl Pearson correlation coefficient (CC), denoted by r between x and y:

$$r = \frac{n\sum xy-(\sum x)(\sum x)}{\sqrt{n\sum x^2-(\sum x)^2}\sqrt{n\sum y^2-(\sum y)^2}} \quad (1)$$

Properties of Correlation Coefficient:

a) CC always lies between -1 and +1.
b) When CC equals to +1, there is perfect positive linear relationship.
c) When CC equals to -1, there is perfect negative linear relationship.
d) When CC equals to 0, there is no linear relationship.
e) When CC is positive, there is positive correlation.
f) When CC is negative, there is negative correlation.
g) When CC is near to -1 or $+1$, there are strong linear relations.

Relative Error (RE):

The RE reflects the credibility of the proposed method, and is calculated by the formula [Qiang Fu, et al. (2018)]:

$$RE = \frac{|y-x|}{x} \times 100\ \% \quad (2)$$

Scatter Diagram:

Scatter diagram also called scatter plot, dot diagram, is just a approach to study the relationship between two variables. When the pair of values (x_i, y_i) for i=1, 2, …, n are plotted on a two-dimensional plane, the points show the pattern in which they actually lie. Such a diagram or pattern is known as scatter diagram. If these points lie on a straight line, it is expected that there is a linear relationship between x and y, otherwise not. It is a pictorial representation of the data, which demonstrates us the direction of the relationship between two variables.

Use a scatter diagram technique to examine theories about cause-and-effect relationships and to search for root causes of an identified problem. Use a scatter diagram to design a control system to ensure that gains from quality improvement efforts are maintained.

Linear Regression:

Linear regression is basically a statistical process that attempts to model the relationship between two variables. This is carried out by fitting a linear equation to observed data. One variable is treated as an explanatory variable, and the other is treated as a dependent variable.

Before start to attempt to fit a linear regression model to the observed data, a decision maker (DM) must first determine whether or not there is a relationship between the two variables. This does not necessarily imply that one variable causes the other, but that there is some significant association between the variables of interest.

The equation of a line of linear regression of the form

$$y = a\,x + b \quad (3)$$

where **x** is the independent variable and is the dependent variable. The slope of line is $\boldsymbol{a}$, and $\boldsymbol{b}$ is the value of **y** when **x** = 0.

III. NOTATIONS

In this manuscript, the following notations are adopted:

x:Time from Submission to first decision, in days.
y: Acceptance rate, in %.
n: Number of data points.
CC:Karl Pearson correlation coefficient (r). RE : Relative error.
$\overline{RE}$: Mean relative error.
$\sum x$: Sum of all x terms.
$\sum y$: Sum of all y terms.

IV. PROBLEM DEFINITION

Our objective is to determine: how the acceptance rate (y) in percentage is related with the first decision time (x) in days. For this purpose, we plot a scatter diagram to know the regression is linear or non-linear. Also, we calculate the Karl Pearson correlation coefficient (r).

We collected the real data for 39 indexed journals dated 31-03-2019, as given below in Table 1, from the following web sources:

Web Source (i) Elsevier Journal Finder, link:*https://journalfinder.elsevier.com*
Web Source (ii) Springer Journal Suggester/Finder, link:*https://journalsuggester.springer.com*

Table 1: Real Date Collected from Sources (i) & (ii)

Journal	Abstracts & Indexing	Time from Submission to First decision,	Time from Submission to first decision, (x) in days	Acceptance Rate (AR) In % (y)
Applied Mathematics Letters	SCOPUS, ESCI	1 week	7	11
Decision Support Systems	SCOPUS, ESCI, SCI	3 weeks	21	11
Information Processing & Management	SCOPUS, ESCI, SCI	3 weeks	21	9
Transportation Research Part E: Logistics & Transportation Review	SCOPUS, SCI	4 weeks	28	15
Expert Systems with Applications	SCOPUS, SCIE	5 weeks	35	13
Knowledge-Based Systems	SCOPUS, SCI,	6 weeks	42	17

Operations Research Letters	SCOPUS, SCIE	8 weeks	56	26
Int. J. of Approximate Reasoning	SCOPUS, ESCI, SCI	8 weeks	56	30
Neurocomputing	SCOPUS, ESCI	9 weeks	63	29
Computers & Operations Research	SCOPUS, SCI	9 weeks	63	14
Applied Mathematical Modelling	SCOPUS, SCI	9 weeks	63	17
Applied Soft Computing	SCOPUS, SCI	8 weeks	56	16
Information Science	SCOPUS, SCI,	8 weeks	56	19
Omega	SCOPUS, SCI	6 weeks	42	12
Information Fusion	SCOPUS, SCI	8 weeks	56	19
Journal of Manufacturing Systems	SCOPUS, SCI	6 weeks	42	15
Journal of The Franklin Institute	SCOPUS, SCI	11 weeks	77	31
Applied Energy	SCOPUS, SCI	5 weeks	35	14
Operations Research Perspectives	SCOPUS, SCI	5 weeks	35	13
Int. J. on Interactive Design and Manufacturing	SCOPUS, ESCI	16 days	16	39
Palgrave Communications	SCOPUS, ESCI	40 days	40	38
Applied Intelligence	SCOPUS, SCI-E	19 days	19	19
Computational Optimization and Applications	SCOPUS, SCI, SCI-E	45 days	45	14
J. of Global Optimization	SCOPUS, SCI, SCI-E	60 days	60	24
EURO Journal on Computational Optimization	SCOPUS, SCI	67 days	67	27
Optimization Letters	SCOPUS, SCIE	77 days	77	26
Japan Journal of Industrial and Applied Mathematics	SCOPUS, SCI, SCI-E	86 days	86	32
Int. J. of System Assurance Engineering and Management	SCOPUS, SCIE	127 days	127	47
J. of Industrial Engineering International	SCOPUS	109 days	109	23
Int. J. of Applied & Computational Mathematics	SCOPUS	76 days	76	38
Int. Journal of Fuzzy Systems	SCOPUS, SCI-E	48 days	48	22
J. of the Operations Research Society of China	SCOPUS, ESCI	99 days	99	43
Queueing Systems	SCOPUS, SCI,	95 days	95	44
Int. Journal of Dynamics & Control	SCOPUS	60 days	60	35
J. of Intelligent Manufacturing	SCOPUS, SCI-E	32 days	32	12
J. of Optimization Theory and Applications	SCOPUS, SCI, SCI-E,	60 days	60	18
Annals of Data Science	Non-ESCI. Non-SCOPUS	61 days	61	63
Granular Computing	Non-SCOPUS,	30 days	30	42
Bulletin of the Iranian Mathematical Society	SCOPUS, SCIE	20 days	20	32

V. NUMERICAL CALCULATIONS

We take the data from the Table 1, for the variable x as well as y. To determine the correlation coefficient (CC), r, and RE, we prepare the following Table 2 as below:

Table 2: Data for the Determination CC, and RE.

x	y	xy	x^2	y^2	RE= $\frac{\lvert y-x\rvert}{x}\times 100$ %
7	11	77	49	121	57.14
21	11	231	441	121	47.61
21	9	189	441	81	57.14
28	15	420	784	225	46.42
35	13	455	1225	169	62.85
42	17	714	1764	289	59.52
56	26	1456	3136	676	53.57
56	30	1680	3136	900	46.42
63	29	1827	3969	841	34.00
63	14	882	3969	196	49.00
63	17	1071	3969	289	73.01
56	16	896	3136	256	71.42
56	19	1064	3136	361	66.07
42	12	504	1764	144	71.42
56	19	1064	3136	361	66.07
42	15	630	1764	225	64.28
77	31	2387	5929	961	59.74
35	14	490	1225	196	60.00
35	13	455	1225	169	62.85
16	39	624	256	1521	143.75
40	38	1520	1600	1444	5.00
19	19	361	361	361	0
45	14	630	2025	196	68.88
60	24	1440	3600	576	60.00
67	27	1809	4489	729	59.70
77	26	2002	5929	676	66.23
86	32	2752	7396	1024	63.79
127	47	5969	16129	2209	62.99
109	23	2507	11881	529	78.89
76	38	2888	5776	1444	50.00
48	22	1056	2304	484	54.16
99	43	4257	9801	1849	56.56
95	44	4180	9025	1936	51.00
60	35	2100	3600	1225	41.66
32	12	384	1024	144	62.50
60	18	1080	3600	324	70.00
61	63	3843	3721	3969	3.27
30	42	1260	900	1764	40.00
20	32	640	400	1024	60.00
$\sum x$=2081	$\sum y$=969	$\sum xy$=57794	$\sum x^2$=138015	$\sum y^2$=30009	$\sum RE$=2206.91

Substituting all the summation values from Table 2 in equation (1), the value of correlation coefficient r can be determined. The value of r is:

$$r = +0.4813 \qquad (4)$$

Mean value of RE is given by

$$\overline{RE} = \frac{\sum RE}{n} = \frac{2206.91}{39} = 56.58\ \% \qquad (5)$$

which shows that there is a positive correlation between x (time from submission to first decision, in days) and y (acceptance rate, in %). The strength of this correlation is not so strong, and not so weak, rather is moderate correlation. The scatter diagram is also platted as given in the following Figure 1.

Calculation of Equation of Linear Regression:

$\sum x$=Sum of x = 2081

$\sum y$ =Sum of y = 969

Mean X = 53.359

Mean Y = 24.8462

Sum of squares (SS_X) = 26974.9744

Sum of products (SP) =6089.1538

Regression Equation:

$y = bx + a$

where $b = SP/SS_X$

$= 6089.15/26974.97$

$= 0.22573$

$a = M_Y - bM_X$

$= 24.85 - (0.23*53.36)$

$= 12.80125$

$y = 0.22573x + 12.80125 \qquad (6)$

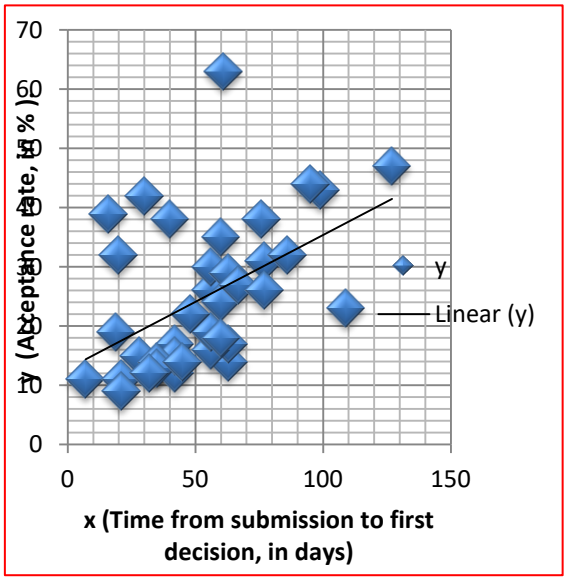

Figure 1: Scatter Diagram

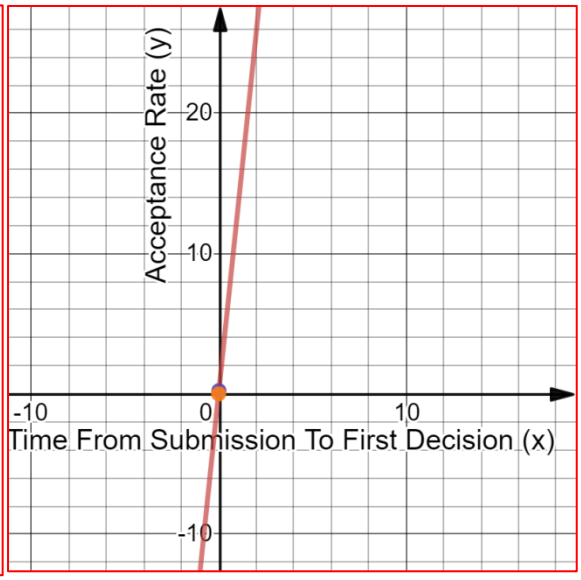

Figure 2: Straight Line Graph

Particular Cases

In this section, we discuss the following two cases:

Case I: Determination of y, when x=10

By equation (6), we obtain

$y = 0.22573(10) + 12.80125$

$= 13 \qquad (7)$

When the time from submission to first decision is 10 days, the acceptance rate would be 13.00 %.

Case II: Determination of y, when x=300

By equation (6), we obtain

$y = 0.22573(300) + 12.80125$

$= 80.52 \qquad (8)$

When the time from submission to first decision is 300 days, the acceptance rate would be 80.52 %.

Table 3: Experimental Results

	Time from submission to first decision (days)	Acceptance rate
Case I	**10**	**13.00 %**
Case II	**300**	**80.52 %**

VI. CONCLUSIONS

In the present study, we determined a relationship between the first decision time, in days, and the acceptance rate, in percentages, for some indexed journals, from Elsevier and Springer publishers. This study is based on the real data, collected from the web sources of these journal publishers. The value of correlation coefficient is $+0.4813$, which represents a positive correlation, whose strength is moderate. The mean relative error ($\overline{RE}$) is calculated to support the proposed technique. Additionally, we analyzed the regression, which we found in linear trend. A linear relationship between these measures of the various journals is concluded. In this manuscript, a lot of information is given about the top indexed journals. The significant contribution of this paper may be very useful to the researchers and academicians in choosing a suitable journal as per acceptance rate and first decision time.

Due to manuscript drafting only with correlation and regression analysis, some important aspects remain untouched and may be explored as a future research scope. For example, the fuzzy and probabilistic nature of these measures may also be another dimension for further research work.

The proposed approach could be further extended by incorporating the multi-choice and fuzzy stochastic parameters along with multi-level decision-making scenarios.

FUNDING

This research received no external funding. Hence, the founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

ACKNOWLEDGMENTS

The author is deeply thankful to potential reviewers and editors for their valuable comments and suggestions to improve the quality and readability of the paper.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

REFERENCES

1. Zeger, S.L., (1988). A regression model for time series of counts. Biometrika, 75, 621-629.
2. Hair, J.F., Anderson, R.E., Tatham, R.L., & Black, W.C. (1998). Multivariate data analysis. New Jersey: Prentice-Hall International, Inc.
3. Natalie, M., & James, R. (2002). An improved batch means procedure for simulation output analysis, Management Science, 48(12), 1569-1586.
4. Koenker, R. (2004). Quantile regression for longitudinal data, Journal of Multivariate Analysis, 91(1), 74-89.
5. Canay, I.A. (2011). A simple approach to quantile regression for panel data, *Econometrics Journal*, 14 (3), 368-386.
6. Azagba, S., Sharaf, M. F. (2012). Fruit and vegetable consumption and body mass index: a quantile regression approach, Journal of Primary Care & Community Health, 3(3), 210–220.
7. Wang, J. (2012). Bayesian quantile regression for parametric nonlinear mixed effects models, Statistical Methods & Applns, 21(3), 279-295.
8. Ismail, Azman; Ahmad Azan Ridzuan; Nur Ilyani Ranlan Rose; Muhammad Madi Bin Abdullah; Muhammad Sabbir Rahman; Francis, & Sebastian K., (2013). Examining the relationship between service quality and customer satisfaction: A factor specific approach, Journal of Industrial Engineering and Management, 6(2), 654-667.
9. Osman, Z., & Sentosa, I. (2013). Mediating effect of customer satisfaction on service quality and customer loyalty relationship in Malaysia Rural Tourism. International Journal of Economics Business and Management Studies, 2(1), 25-37.
10. Qiu, Heting; Li, & Xuewei, (2013). A study of combined evaluation of suppliers based on correlation, Journal of Industrial Engineering and Management, 6(1), 249-257.
11. Bottai, M., Frongillo, E. A., & Sui, X. (2014). Use of quantile regression to investigate the longitudinal association between physical activity & body mass index, Obesity, 22(5), E149–E156.
12. Wang, Yin-ying; Lin, Shang-ping; Fang, Ching-han; Lo, & Shi-hwa, (2015). The canonical correlation analysis between the fitness of organizational condition and integration possibility: The evidence of configurational approach, International Journal of Management, Economics and Social Sciences, 4(3), 92-107.
13. Geraci, M. (2016). Estimation of regression quantiles in complex surveys with data missing at random: An application to birthweight determinants, Statistical Methods in Medical Research, 25(4), 1393-1421.
14. Coulom, Jean Shenai, & Vijay, (2018). The effect of alternative measures of distance on the correlation of real effective exchange rate returns: An approach to contagion analysis, International Journal of Financial Studies, 6(4), 1-18.
15. Kumar, P., Keerthika, P. S., (2018). An inventory model with variable holding cost and partial backlogging under interval uncertainty: Global criteria method. International Journal of Mechanical Engineering and Technology, 9(11), 1567-1578.
16. Prameela, K.U., Kumar, P., (2019). Execution proportions of multi server queuing model with pentagonal fuzzy number: DSW algorithm approach. International Journal of Innovative Technology and Exploring Engineering, 8(7), 1047-1051.
17. Web source Elsevier Journal Finder, link: *https://journalfinder.elsevier.com*
18. Web source Springer Journal Suggester/Finder, link: *https://journalsuggester.springer.com*

AUTHOR PROFILE

PAVAN KUMAR is an Assistant Professor of Mathematics at the Koneru Lakshmaiah Education Foundation (K L University), Vaddeswaram (AP), India. He received his PhD (2014) in Operations Research, from NIT Warangal; NET−JRF (CSIR−UGC, 2009) qualified. He has published several research papers in international, national journals of high repute. His main research interests are mathematical programming, fuzzy optimization, stochastic optimization, inventory control problems, transportation problems, and other applications of mathematical programming. His e-mail address is <pavankmaths@gmail.com>.

ORCID Dr.Pavan Kumar

https://orcid.org/0000-0001-5340-7777

Dr. W.Sridhar working as Associate professor in the Department of Mathematics, Koneru Lakshmaiah Education Foundation Deemed to be University, Vaddeswaram, Guntur dist, A.P., He has 16 Publications in various national and International Journals. He did his Ph.D from Adikavi Nannaya University, Rajamahendravaram, A.P, India. His current area of research on Flow ,Heat and Mass transfer analysis of Viscous fluids.

ORCID Dr.W. Sridhar

https://orcid.org/0000-0002-0978-1769